Explaining the persistent influence of facial cues in social decision-making

Jaeger, Bastian; Evans, Anthony; Stel, M.; van Beest, Ilja

Published in:
Journal of Experimental Psychology: General

DOI:
10.1037/xge0000591

Publication date:
2019

Document Version
Peer reviewed version

Explaining the persistent influence of facial cues in social decision-making

Bastian Jaeger and Anthony M. Evans
Tilburg University
Mariëlle Stel
Twente University
Ilja van Beest
Tilburg University

This manuscript has been accepted for publication at the *Journal of Experimental Psychology: General*. Feedback and comments are welcome.

Author Note

Bastian Jaeger, Anthony M. Evans, and Ilja van Beest, Department of Social Psychology, Tilburg University, The Netherlands; Mariëlle Stel, Department of Psychology of Conflict, Risk and Safety, The Netherlands.

Parts of the data reported here were presented at the 2016 TIBER symposium, the 2016 ASPO conference, the 2016 SJDM meeting, and the 2017 ICSD meeting. A pre-print of this paper was shared on PsyArxiv. We thank Patrick Heck and Joachim Krueger for their valuable comments.

Correspondence concerning this article should be addressed to Bastian Jaeger, Department of Social Psychology, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands. E-mail: bxjaeger@gmail.com.

Abstract

Impressions of trustworthiness based on facial cues influence many consequential decisions, in spite of their (generally) poor accuracy. Here, we test whether reliance on facial cues can be better explained by (a) the belief that facial cues are more valid than other cues or by (b) the quick and primary processing of faces, which makes relying on facial cues relatively effortless. Six studies (N = 2,732 with 73,182 trust decisions) test the two accounts by comparing the effects of facial cues and economic payoffs on trust decisions. People believe that facial cues are less valid than economic payoffs (Study 1), but relying on facial cues takes less time than relying on economic payoffs (Study 2). Critically, introducing facial cues causes people to discount payoff information, but introducing payoff information does not reduce the effect of facial cues (Studies 3a-c). Finally, when making intuitive (vs. reflective) trust decisions, people rely less on payoff information, but they do not rely less on facial cues (Study 4). Together, these findings suggest that persistent reliance on facial trustworthiness is better explained by the intuitive accessibility of facial cues, rather than beliefs that facial cues are particularly valid.

Keywords: trust, trustworthiness, face perception, judgment and decision-making

Explaining the persistent influence of facial appearance in social decision-making

First impressions are often based solely on a person's facial appearance (Oosterhof & Todorov, 2008). A number of studies have demonstrated that these face judgments, especially judgments of trustworthiness, influence decisions in domains such as voting, personnel selection, and criminal sentencing (Olivola, Funk, & Todorov, 2014; Olivola & Todorov, 2010a; Todorov, Olivola, Dotsch, & Mende-Siedlecki, 2015). The widespread influence of facial trustworthiness is surprising, given that the human ability to accurately identify trustworthy individuals based on facial cues is generally poor (Todorov, Olivola, et al., 2015).

Why do people rely on first impressions based on facial cues? On the one hand, models of cue selection propose that the perceived diagnosticity of a cue (i.e., how well it is thought to predict a certain outcome) determines how much people rely on it when making decisions (Brunswik, 1956; Gigerenzer & Goldstein, 1996; Hammond, Hursch, & Todd, 1964). Consequently, the influence of facial cues has been attributed to the *perceived* validity of face judgments (Olivola et al., 2014; Rezlescu, Duchaine, Olivola, & Chater, 2012): This view suggests that people rely on facial trustworthiness judgments because they *believe* that their judgments are accurate. On the other hand, it has also been suggested that face judgments affect decisions because of their intuitive accessibility (Olivola & Todorov, 2010a; Willis & Todorov, 2006). Faces are processed quickly and efficiently (Freeman & Johnson, 2016) and people tend to prioritize cues that come to mind easily (Evans & Krueger, 2016; Shah, 2007; Simmons & Nelson, 2006). Relying on facial cues might therefore constitute a mental shortcut that reduces decision effort (Gigerenzer, Hertwig, & Pachur, 2011; Shah & Oppenheimer, 2008). Here, we test these two accounts, illuminating the mechanisms that give rise to the widespread influence of facial cues.

Facial trustworthiness influences social decision-making

The ability to judge trustworthiness in others is crucial in mixed-motive settings, where there is a motivational conflict between self-interest and the collective good (Kelley et al., 2003). When this conflict arises, people must judge whether their interaction partners can be trusted to cooperate (Dawes, 1980). Researchers employ simplified interactions, such as the trust game (Figure 1), to capture the essential structure of this dilemma (Berg, Dickhaut, & McCabe, 1995). In the binary version of the game (Snijders & Keren, 1999), the *trustor* can decide to keep the status quo, which ends the interaction, or to trust the other player. In the event of trust, the *trustee*

can choose between betrayal and reciprocation. Reciprocity leads to equal payoffs for both, and these outcomes are better than the status quo. The trustee gains even more by choosing betrayal, but this leads to the worst possible outcome for the trustor. Trust is risky; once it is chosen, the trustee has full control over the outcomes, and is faced with the temptation to choose betrayal.

In the trust game, people rely on facial cues when making decisions (Ewing, Caulfield, Read, & Rhodes, 2014; Rezlescu et al., 2012; van't Wout & Sanfey, 2008). In general, trustworthy-looking partners are more likely to be trusted. Similar effects have been observed for consequential trust decisions outside the lab: Trustworthy-looking individuals ask for higher rent on Airbnb (Ert, Fleischer, & Magen, 2016); have a higher chance of being granted loans on crowdsourcing websites (Duarte, Siegel, & Young, 2012); and are more likely to be appointed as CEOs after firm misconduct (Gomulya, Wong, Ormiston, & Boeker, 2017). Facial trustworthiness also influences legal decisions (Porter, ten Brinke, & Gustaw, 2010; see also Berry & Zebrowitz-McArthur, 1988). Wilson and Rule (2015, 2016) found that untrustworthy-looking criminals were more likely to receive the death penalty (as opposed to life in prison). In short, perceived facial trustworthiness has far-reaching consequences.

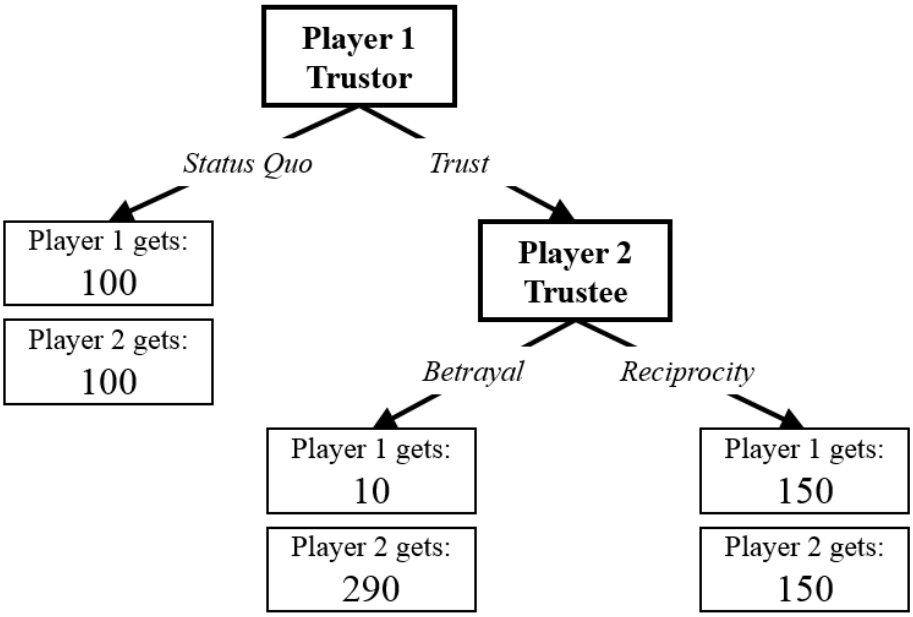

Figure 1. The sequential trust game with exemplary payoffs.

Despite widespread reliance on facial cues, evidence on whether people are able to accurately infer others' trustworthiness from their facial appearance is mixed (Bonnefon,

Hopfensitz, & De Neys, 2015; Todorov, Funk, & Olivola, 2015). Some studies point to modest accuracy (Bonnefon, Hopfensitz, & De Neys, 2013; De Neys, Hopfensitz, & Bonnefon, 2017; C. Lin, Adolphs, & Alvarez, 2018; Slepian & Ames, 2015; Tognetti, Berticat, Raymond, & Faurie, 2013), whereas others provide evidence against it both on empirical (Efferson & Vogt, 2013; Rule, Krendl, Ivcevic, & Ambady, 2013) and theoretical grounds (McCullough & Reed, 2016; Vogt, Efferson, & Fehr, 2013). For example, Todorov and Porter (2014) found substantial variation in judgments of trustworthiness across different photos of the same person, which speaks against the idea that there are stable cues signaling an underlying trait (see also Sutherland, Young, & Rhodes, 2016). More evidence is needed to determine when people are able to accurately discern trustworthiness from facial cues. However, the current literature suggests that this ability is limited at best, and people would often make better decisions by relying on other cues (Bonnefon, Hopfensitz, & De Neys, 2017; Olivola & Todorov, 2010b).

The observation that facial trustworthiness judgments affect so many consequential decisions (in spite of their poor accuracy) has led several researchers to propose that the influence of face judgments constitutes a bias that should be eliminated (e.g., Olivola, Funk, & Todorov, 2014; Porter et al., 2010; Wilson & Rule, 2015). However, the origin of such a face bias (i.e., *why* there are such persistent effects of facial trustworthiness) remains poorly understood, and this shortcoming has hindered efforts aimed at curbing the bias. Here, we set out to address this gap.

Explanations for reliance on facial trustworthiness

Scholars have long been interested in how people use different types of information, or cues, when making decisions (Brunswik, 1956; Slovic & Lichtenstein, 1971). Normative models propose that rational decision-makers should weigh all available cues according to how strongly they correlate with an outcome (Dawes, Faust, & Meehl, 1989). Yet, people often lack insight into the true validity of a cue (e.g., how diagnostic facial trustworthiness judgments really are) and descriptive decision models acknowledge that cues are weighed (or ranked) according to their *subjective*, rather than objective, validity (Brunswik, 1956; Gigerenzer & Goldstein, 1996; Hammond et al., 1964). A weaker definition of rationality holds that cue utilization should be determined by how valid people *think* the cue is. Reliance on facial trustworthiness may then be explained by the subjective belief that it is a particularly valid cue. This argument has been raised in a review by Olivola and colleagues (2014), who point to the historically persistent

belief in the correspondence between facial cues and personality (Aristotle, trans. 1936; Lavater, 1775). For example, Hassin and Trope (2000) report that most people believe that faces contain at least some valid cues to individual personality traits (see also Suzuki, Tsukamoto, & Takahashi, 2017). We refer to this explanation of the face bias as the *subjective validity account*.

An alternative explanation can be derived from research into heuristic decision-making. People are often unable or unwilling to consider all available information and therefore rely on heuristics—strategies that consider only a subset of all available information—to reduce cognitive effort (Gigerenzer et al., 2011; Payne, Bettman, & Johnson, 1988; Simon, 1955). In fact, people favor cues that are intuitively available (Dimov & Link, 2017; Shah, 2007; Simmons & Nelson, 2006). Faces attract attention (Ro, Russell, & Lavie, 2001; Theeuwes & Van der Stigchel, 2006) and trustworthiness impressions are formed spontaneously (Klapper, Dotsch, van Rooij, & Wigboldus, 2016), quickly (Willis & Todorov, 2006), and effortlessly (Bonnefon et al., 2013). From this perspective, reliance on facial trustworthiness may be explained by the intuitive accessibility of the cue (Olivola & Todorov, 2010a; Willis & Todorov, 2006). We refer to this explanation of the face bias as the *intuitive accessibility account*.

Overview of the studies

We present six studies to test the two accounts of the face bias: Do people persistently rely on facial trustworthiness because they believe it is a particularly valid cue, or because face judgments are intuitively accessible? We examine this question in the context of the trust game (Figure 1, Berg et al., 1995; Snijders & Keren, 1999), by comparing the effects of facial trustworthiness and economic payoff information, a cue we predict to be more subjectively valid, but also more difficult to evaluate. The combination of these two cues, which differ in subjective validity and accessibility, offers a critical test of the factors underlying cue preferences.

In Study 1, we test the perceived validities of facial trustworthiness and economic payoff information, with the prediction that people believe economic payoffs are more valid cue than facial trustworthiness. In Study 2, we examine the time it takes to make decisions based on facial trustworthiness and economic payoffs; here, we expect it takes less time to reach a decision based on facial trustworthiness than on economic payoffs. Studies 3a-c examine how much people rely on the two cues when both are presented simultaneously. In these studies, we determine whether people favor the subjectively more valid cue (economic payoffs), or the cue that is processed more efficiently (facial trustworthiness). Finally, Study 4 tests how cognitive

reflection affects reliance on the two cues. If people favor facial trustworthiness because it is easier to process than payoff information, then making intuitive (vs. reflective) decisions should reduce reliance on economic payoffs, but not reduce reliance on facial trustworthiness.

All data and analysis scripts are available at the Open Science Framework (https://osf.io/h6dsj/).

Study 1

Study 1 examined explicit preferences for facial trustworthiness and economic payoff information. Evans and Krueger (2011, 2014) found that trustors are aware of the importance of the trustee's economic payoffs—people trust less when the trustee faces a greater temptation (i.e., a greater economic incentive to choose betrayal). In turn, trustees actually reciprocate less when temptation is large (Evans & Krueger, 2014). Thus, like facial trustworthiness, temptation is used as a cue when making trust decisions. However, prior work has not examined which of the two cues is seen as more valid. Given that temptation is actually predictive of trustees' behavior while facial trustworthiness has poor predictive validity, we set out to test whether people's explicit preferences for the two cues correspond to their predictive validities. To address this question, we let participants play a trust game where participants could choose which cue they would want to have available: temptation (the trustee's economic payoff to betray trust) or facial trustworthiness (a facial photograph of their interaction partner).

Method

Participants. Dutch undergraduate psychology students (N = 126; 81.0% female) from Tilburg University participated in the study in return for partial course credit (M_{age} = 19.83, SD_{age} = 2.96). The sample size was based on the number of students that participated in the study within one week.

Materials and procedure. The experiment was administered in the lab. First, participants saw an exemplary decision tree (see Figure 1) and learned about the rules of the trust game. They were told that they would have to make a single decision between IN and OUT (trust and status quo) and that they would be randomly paired with another participant who would act as the trustee. If they chose OUT, the interaction would end. If they chose IN, then the trustee would make a choice between RIGHT and LEFT (reciprocation and betrayal). Choices were hypothetical and participants were asked to imagine that the points they were playing for would

be converted into actual money at the end of the game. All participants played the role of the trustor.

Next, participants chose what kind of information (i.e., economic payoff or facial appearance) they wanted to have available. They could choose to see *how much your interaction partner would gain in case he/she chooses LEFT* or *a photo of your interaction partner*. The order in which they read the description of the two cues was counterbalanced. Participants then saw a different decision tree (with different payoffs) that was relevant for their choice. Depending on their choice of cue, either the betrayal payoff or the photo of their partner was revealed. The photo was taken from the Radboud Faces Database (RaFD; Langner et al., 2010) and showed a female person with a happy facial expression. The trustee's temptation corresponded to a 100% increase in payoffs when choosing betrayal over reciprocation. Participants made their one-shot trust decision and indicated their confidence in having made the right decisions by dragging a slider along a scale ranging from 0 (*not at all confident*) to 100 (*extremely confident*).

Results and discussion

To measure preferences for the two types of cues, we compared the percentages of participants who chose to see the photo of their partner versus the payoffs of their partner. A clear majority of participants (78.6%) preferred to see the trustee's temptation, $p < .001$ (exact binomial test). Participants explicitly prioritized economic payoff information over facial trustworthiness information.

After cue selection, a total of 23.02% of participants trusted. Confidence ratings were similar for trust ($M = 58.45$, $SD = 18.85$) and distrust decisions ($M = 62.92$, $SD = 23.11$), $t(55.54) = 1.06$, $p = .29$, $d = 0.22$.[1] There was also no significant difference in confidence ratings between participants who chose to see their partner's payoffs ($M = 61.86$, $SD = 22.80$) and those who chose to see their partner's photo ($M = 62.00$, $SD = 20.30$), $t(45.53) = 0.03$, $p = .98$, $d = 0.01$. Confidence was not affected by the type of information available at the moment of decision-making.

[1] We report the results of Welch's *t*-tests which, compared to a Student's *t*-tests, provides equal power but superior error control in case of unequal variances between groups (Delacre, Lakens, & Leys, 2017).

Study 2

In our first study, people perceived economic payoff information to be more valid than facial trustworthiness information. Our second study had two goals: First, we wanted to compare the cognitive effort required to make decisions based on payoffs versus facial trustworthiness. Previous research has shown that facial trustworthiness is processed spontaneously, quickly, and effortlessly (Bonnefon et al., 2013; Klapper et al., 2016; Willis & Todorov, 2006). However, it is still unclear if making decisions based on facial trustworthiness is less effortful than making decisions based on other cues. Reduced decision effort should be reflected in faster decision times (cf. Bettman, Johnson, & Payne, 1990). We therefore hypothesized that people would make decisions more quickly when they relied on facial trustworthiness rather than economic payoffs. Differences in response times may also be caused by differences in decision conflict rather than decision effort (Evans & Rand, 2019; Krajbich, Bartling, Hare, & Fehr, 2015). Thus, we measured participants' confidence, as increased decision conflict leads to decreased confidence (De Neys, Cromheeke, & Osman, 2011; Zakay, 1985).

Second, we examined how facial and economic cues influenced the extent to which people rely on expectations of reciprocity. Expectations play a central role in psychological and economic models of trust, providing a conceptual link between cues and trust decisions (Bacharach & Gambetta, 2001; Thielmann & Hilbig, 2015). People are more likely to trust if they have high expectations of reciprocity (Costa-Gomes, Huck, & Weizsäcker, 2014; Thielmann & Hilbig, 2014). The subjective validity account suggests that people should rely more on their expectations when they are based on economic payoffs (vs. facial trustworthiness) since economic payoffs are subjectively more valid. The intuitive accessibility account makes the opposite prediction, as using appearance to form expectations requires less cognitive effort. Thus, comparing the extent to which people rely on their cue-based expectations served as a critical test of the two accounts.

Method

Participants. A total of 134 students from Tilburg university (75.4% female; M_{age} = 21.30, SD_{age} = 1.45) participated in exchange for partial course credit. The sample size was based on the number of students that participated in the study within two weeks.

Materials and procedure. The experiment was administered online. Participants were randomly assigned to the temptation condition or the face condition. In both conditions,

participants first learned about and then played a series of 24 hypothetical trust games in the role of the trustor.

In the face condition, participants saw a photo of their interaction partner next to the decision tree on each trial. The photos were again taken from the RaFD (Langner et al., 2010). We selected 24 frontal photos of Caucasian Dutch adults with a forward gaze, of which half were male and half were female. Similar to previous investigations, half of the selected faces displayed a neutral expression and half a happy facial expression (i.e., they were smiling) in order to introduce variance in the perceived trustworthiness of the faces (cf. Evans & van de Calseyde, 2017). Previous research has shown that smiling individuals are perceived to be more trustworthy (Krumhuber et al., 2007; Said, Sebe, & Todorov, 2009). To ensure that any observed effect of facial trustworthiness is not due to the attractiveness of the face (R. K. Wilson & Eckel, 2006), we selected faces of 12 male and 12 female individuals judged to be equally attractive (Langner et al., 2010).

In the temptation condition, we varied the trustee's temptation, i.e., the economic incentive to choose betrayal. Following previous work by Evans and Krueger (2014), we defined temptation as the difference between the trustee's gain in case of betrayal (T) and reciprocation (R_2) divided by the value of betrayal: $(T - R_2)/T$ (see Figure 1). On half of the trials, temptation was low (0.33) and on the other half, it was high (0.60). These values correspond to a 50% (low temptation) and 150% (high temptation) increase in payoffs for the trustee in case betrayal is chosen over reciprocation. Note that payoffs in the face condition always corresponded to a temptation parameter of 0.5 (100% increase in payoffs) which ensured that the average temptation across the 24 trials was equal in both conditions.

We assessed participants' response time for each trust game decision. Five extremely slow responses (0.16% of all decisions) were excluded because they were between 3 and 56 standard deviations slower than the mean. Response times were $\log_{10}$-transformed to account for their right-skewed distribution.

After making decisions in the 24 trust games, participants were shown each trust game again and they were asked to state their expectations of reciprocity. Specifically, we asked: *How likely is it that Player 2 will choose RIGHT* [reciprocation]*?* They could drag a slider along a scale ranging from 0 (*Player 2 will definitely choose LEFT* [betrayal]) to 100 (*Player 2 will definitely choose RIGHT* [reciprocation]). Participants also rated how confident they are that

their expectations are accurate by dragging a slider along a scale ranging from 0 (*I am not confident at all*) to 100 (*I am extremely confident*). Following this, participants in the face condition saw each face one more time and were asked to rate how trustworthy the person in the photo is. Again, they could drag a slider along a scale ranging from 0 (*not trustworthy at all*) to 100 (*extremely trustworthy*). Similar to previous work, we used the average trustworthiness rating of each face across all participants as our measure of facial trustworthiness (van't Wout & Sanfey, 2008).[2]

Results

Descriptive statistics. Average trustworthiness ratings of the faces ranged from 33.80 to 69.45 (M = 51.10 out of 100, SD = 11.78) and participants showed significant consensus in their ratings of the faces, ICC = .314, $p < .001$, 95% CI [.204, .499]. Overall, participants trusted on 37.47% of all trials and it took them on average 5.53 seconds (SD = 6.69) to reach a decision. Sixteen participants (11.94%) never trusted whereas none of our participants always trusted. Average expectations of reciprocity were below the scale midpoint (50), indicating that participants believed that, across all trials, trustees were more likely to betray than to reciprocate trust (M = 43.23, SD = 16.49), $t(133) = 4.75$, $p < .001$.

Response times. Making the same decision on all trials may lead to faster response times and ultimately obscure how response times are related to cue reliance. Results indeed showed that the sixteen participants who never trusted made significantly faster decisions (M = 0.394, SD = 0.326) than participants whose trust decisions varied across trials (M = 0.602, SD = 0.200), $t(16.57) = 2.49$, $p = .024$, $d = 0.66$. We therefore excluded these participants from our response time analyses.

We compared the time participants took to make decisions based on temptation versus facial trustworthiness. We hypothesized that decisions based on facial trustworthiness would be faster than decisions based on temptation. A t-test comparing response times between the two conditions showed that participants in the face condition reached their decisions substantially faster (M = 0.437, SD = 0.198) than participants in the temptation condition (M = 0.709, SD = 0.167), $t(125.4) = 8.58$, $p < .001$, $d = 1.58$ (see Figure 2A).

[2] In Study 2 and Study 3a, we also measured individual differences in risk-taking, what participants wanted to do, and what they thought they should do for exploratory purposes.

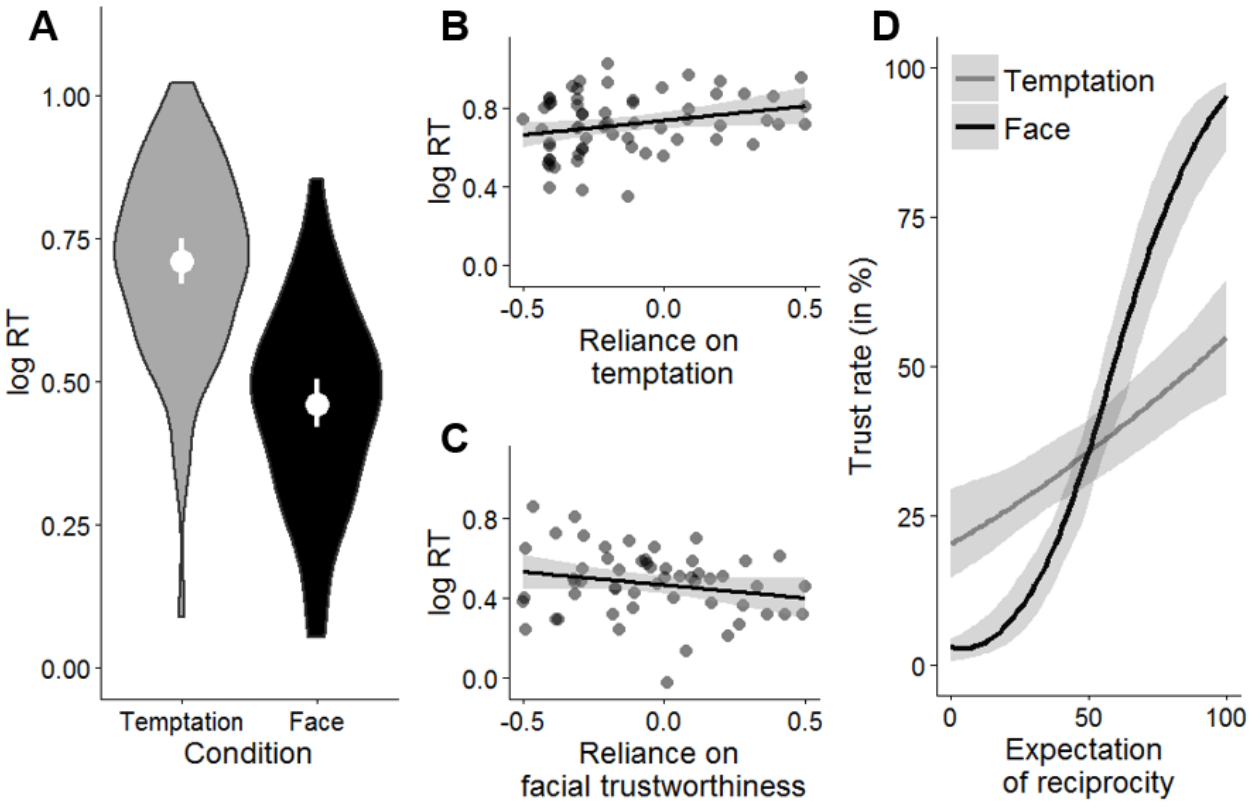

Figure 2. The effects of facial cues and payoff information on response times and reliance on expectations (Study 2): (A) Violin plots showing the difference in $\log_{10}$-transformed response times between the temptation condition and the face condition. Dots denote the mean and bars denote the 95% confidence interval; (B) the correlation between participants' reliance on temptation and their average response times; (C) the correlation between participants' reliance on facial trustworthiness and their average response times; (D) the relationship between expectations of reciprocity and the probability of trust when expectations were based on temptation vs. facial trustworthiness. Values denote the predicted probability of trust derived from multilevel regression models.

We also correlated each participant's average response time with the extent to which they relied on the available cue *within* each condition.[3] To test the hypothesis that reliance on

[3] We extracted the absolute effects of temptation and facial trustworthiness (depending on the condition) on trust decisions from our multilevel regression models as an indication of the extent

temptation is more effortful than reliance on facial cues, we regressed participants' average response times on their level of cue reliance, our condition variable (coded -0.5 for the temptation condition and 0.5 for the face condition), and an interaction term of the two variables. Response times were faster in the face condition, b = -0.272, SE = 0.031, $p < .001$, and cue reliance (across both conditions) had no significant effect on response times, b = 0.008, SE = 0.053, p = .89. More importantly, we found a significant interaction effect between cue reliance and condition, b = -0.278, SE = 0.106, p = .010.

To understand this interaction, we examined the correlation between cue reliance and average response time within each experimental condition. In the temptation condition, we found a positive correlation between the extent to which participants' relied on temptation and the average time needed to make a decision, $r(61) = .265$ p = .036 (see Figure 2B). In the face condition, we found no significant correlation between the extent to which participants relied on facial trustworthiness and their response times, $r(53) = -.214$, p = .12 (see Figure 2C). Increased reliance on temptation was associated with longer response times, providing further evidence that reliance on temptation is more effortful than reliance on facial trustworthiness.

Response time differences between the two conditions may also be driven by differences in decision conflict (Krajbich et al., 2015). To test this account, we compared participants' confidence ratings between the two conditions. If participants made slower decisions based on temptation because they experienced more decision conflict, they should also show decreased confidence (De Neys et al., 2011; Zakay, 1985). On average, participants in the face condition (M = 60.51, SD = 14.38) were not more confident than participants in the temptation condition (M = 64.45, SD = 14.18), $t(131.28) = 1.60$, p = .11, d = 0.28. We also found no evidence that confidence was related to how much people relied on temptation, $r(67) = -.154$, p = .21, or on facial trustworthiness, $r(63) = -.124$, p = .32. Thus, results showed no evidence that response time differences between the two conditions were due to differences in decision conflict.

Reliance on expectations of reciprocity. To conclude, we examined how expectations of reciprocity influenced participants' trust decisions. We estimated multilevel regression models including random intercepts and random slopes. Regressing trust decisions on expectations,

---

to which each participant relied on the cue. Like all other predictors, we standardized this cue reliance variable.

condition, and an interaction term of the two variables yielded a positive effect of expectations, b = 0.038, SE = 0.004, $p < .001$, OR = 1.04, and a negative effect of condition, b = -2.095, SE = 0.450, $p < .001$, OR = 0.12. More importantly, we found a significant interaction effect between expectations and condition, b = 0.042, SE = 0.009, $p < .001$, OR = 1.04.[4] To probe this interaction effect, we examined the effects of expectations within each condition. In line with the intuitive accessibility account, the effect of expectations on trust was larger in the face condition, b = 0.067, SE = 0.010, $p < .001$, OR = 1.07, than in the temptation condition, b = 0.015, SE = 0.004, $p < .001$, OR = 1.02 (see Figure 2D). Thus, participants relied more on their expectations of reciprocity when these could be formed based on the cue that is more intuitively accessible, rather than based on the cue that is seen as more valid.

Discussion

The results of Study 2 shed more light on how facial trustworthiness and economic payoffs differently influence trust decisions. Our results extend previous findings on the processing of facial trustworthiness (e.g., Bonnefon et al., 2013; Klapper et al., 2016; Willis & Todorov, 2006) by showing that making decisions based on facial trustworthiness requires less cognitive effort than making trust decisions based on payoffs. Trust decisions based on payoffs took longer than trust decisions based on facial trustworthiness, and reliance on economic payoffs, but not facial trustworthiness, was positively correlated with decision time.

We also found that the expectations of reciprocity were positively correlated with trusting behavior. Expectations influenced trust decisions when they were based on either payoffs or facial trustworthiness. Crucially, consistent with the intuitive accessibility account, participants relied more on their expectations when they were based on facial trustworthiness (even though Study 1 showed that facial trustworthiness is seen to be a less valid cue). We suggest that this is due to the fact that forming expectations based on easily accessible face judgments is less effortful than considering economic payoffs.

[4] We also examined the possibility of a non-linear relationship between expectations and trust decisions by estimating regression models that included linear, quadratic, and cubic terms for the effect of expectations. A comparison of model fits showed that the linear model provided the best fit. Adding a quadratic term to the model did not significantly increase model fit (p = .17) and neither did adding a cubic term (p = .89).

Study 3a-c

Our next studies were designed to test predictions of the subjective validity and intuitive accessibility accounts more directly. We examined how the presence of facial trustworthiness information affects reliance on economic payoff information (Study 3a). The intuitive accessibility account predicts that people will rely less on economic payoffs when they can also rely on facial trustworthiness, since it takes less effort to rely on the latter. On the other hand, the subjective validity account predicts that how much people rely on payoffs will not depend on whether or not facial trustworthiness is available, since economic information is seen as more valid.

We also examined how the presence of economic payoff information influenced reliance on facial trustworthiness (Study 3b). The intuitive accessibility account predicts that how much people rely on facial trustworthiness will not depend on whether or not economic payoff information is also available. On the other hand, the subjective validity account predicts that people will rely less on facial trustworthiness information when they can also rely on economic payoff information, since the latter is seen as a more valid cue. This setup also allowed us to address an alternative explanation for the discounting of economic information. It is plausible that any cue is discounted if another cue (that is seen as at least somewhat valid) is available as well. Based on this alternative explanation, one would also expect people to discount facial trustworthiness in the presence of economic payoff information.

After conducting these two initial studies, we ran a third pre-registered study that examined both reliance on temptation and facial trustworthiness in an integrated design (Study 3c). We report aggregated results of the three studies since they examined the same hypotheses with almost identical designs using participants from the same source.

Method

Participants. We recruited a total of 2,007 (Study 3a: n = 201, Study 3b: n = 200, Study 3c: n = 1,606) U. S. American workers from Amazon Mechanical Turk (MTurk; Paolacci & Chandler, 2014) who participated in exchange for $1. Participants who failed an attention check at the end of the study or who indicated having only a poor or basic English proficiency (Study 3a: n = 22, Study 3b: n = 18, Study 3c: n = 292) were excluded from analysis leaving a final sample of 1,675 participants (52.37% female, M_{age} = 34.98, SD_{age} = 10.75).

Materials and procedure. The experiments were administered online. Similar to Study 2, all participants learned the rules of the trust game and made a series of 24 (Study 3a and 3b) or 32 (Study 3c) hypothetical trust game decisions. In Study 3a, participants were randomly assigned to the temptation-only condition or the face-and-temptation condition; in Study 3b, participants were randomly assigned to the face-only condition or the face-and-temptation condition; and in Study 3c, participants were randomly assigned to the temptation-only condition, the face-only condition, or the face-and-temptation condition.

In the temptation-only condition (n = 528), trustee's temptation to betray was varied. On half of the trials, temptation was low (0.2) and on the other half, temptation was high (0.60). These values correspond to a 25% (low temptation) and 150% (high temptation) increase in payoffs for the trustee in case betrayal is chosen over reciprocation.

In the face-only condition (n = 525), participants saw a photo of their interaction partner and the level of temptation was held constant. In Study 3b, we selected twelve photos that were already used in Study 1 (six male and six female matched on attractiveness; three individuals with a happy and three with a neutral expression for each gender). In Study 3c, we selected a total of sixteen photos that were already used in Study 1 (eight male and eight female matched on attractiveness; four individuals with a happy and four with a neutral expression for each gender). Participants interacted twice with each individual.

In the face-and-temptation condition (n = 622), we varied both cues orthogonally. In Study 3a, we selected four photos that were already used in Study 1 (two male and two female matched on attractiveness; one individual with a happy and one with a neutral expression for each gender). Each photo was presented six times—three times paired with low and three times paired with high temptation. In Study 3b and Study 3c, each of the selected photos that were also displayed in the face-only condition was presented twice—once paired with low and once paired with high temptation.

Participants were then shown each face again and asked to rate how trustworthy they think the person in the photo is. The average trustworthiness rating of each face across all participants was used as our measure of facial trustworthiness. We removed 285 decisions (0.56% of all decisions) with extremely slow response times because they were between 3 and 78

standard deviations slower than the mean. Response times were $\log_{10}$-transformed to account for their right-skewed distribution.[5]

Results

We rescaled our cue variables to range from -0.5 to 0.5. Thus, for our two cues, a one-unit increase denotes a change from low to high temptation and a change from the lowest average trustworthiness rating to the highest.

Descriptive statistics. Average trustworthiness ratings of the faces ranged, from 35.97 to 74.48 (M = 54.87, SD = 10.92). Participants showed significant consensus in their ratings of the faces in Study 3a, ICC = .281, p < .001, 95% CI [.104, .848], Study 3b, ICC = .294, p < .001, 95% CI [.170, .548], and Study 3c, ICC = .294, p < .001, 95% CI [.185, .500]. The average trust rate across all trials was 39.71% and the average response time was 3.82 seconds (SD = 3.79). A total of 199 participants (11.88%) never trusted whereas 82 participants (4.90%) always trusted.

Reliance on temptation. First, we tested how much participants relied on temptation when it was the only available cue versus when both facial trustworthiness and temptation were available. To this end we estimated a multilevel regression model with random intercepts per participant and random slopes for all predictors, in which we regressed participants' trust decisions on temptation, condition (coded -0.5 for temptation-only and 0.5 for face-and-temptation), and an interaction term between the two variables. This yielded a negative effect of temptation, b = -0.688, SE = 0.047, p < .001, OR = 0.50, and no effect of condition, b = 0.082, SE = 0.119, p = .49, OR = 0.86. Crucially, we observed a significant interaction effect between temptation and condition, b = 0.198, SE = 0.095, p = .036, OR = 1.22 (see Figure 3A). In line with the intuitive accessibility account, but contrary to the subjective validity account, participants relied somewhat less on temptation when they could also rely on facial trustworthiness, b = -0.589, SE = 0.056, p < .001, OR = 0.55, as opposed to when temptation was the only available cue, b = -0.787, SE = 0.076, p < .001, OR = 0.46.

[5] In Study 3a and Study 3b, we again measured participants' expectations of reciprocity and their confidence in their expectations. For the sake of brevity, we report the results in the Supplemental Materials.

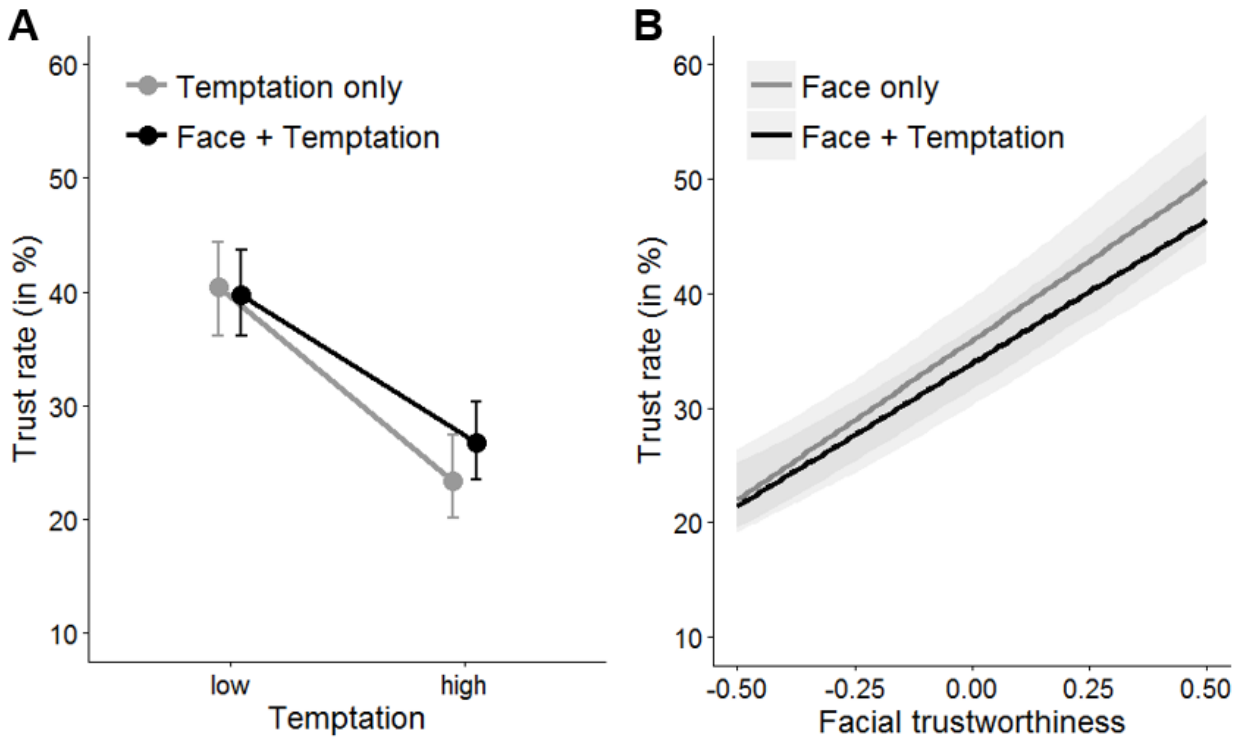

Figure 3. The discounting of economic payoff information in the presence of facial cues (Studies 3a-3c): (A) The effect of temptation on trust rates when temptation was the only available cue vs. when both facial trustworthiness and temptation were available. (B) The effect of facial trustworthiness on trust rates when facial trustworthiness was the only available cue vs. when both facial trustworthiness and temptation were available. Values denote the predicted probability of trust derived from multilevel regression models.

Reliance on facial trustworthiness. Next, we tested how much participants relied on facial trustworthiness when it was the only available cue versus when both facial trustworthiness and temptation were available. To this end, we estimated a multilevel regression model with random intercepts per participant and random slopes for all predictors, in which we regressed participants' trust decisions on the facial trustworthiness of their interaction partner, condition (coded -0.5 for face-only and 0.5 for face-and-temptation), and an interaction term between the two variables. This yielded a positive effect of facial trustworthiness, $b = 1.192$, $SE = 0.071$, $p < .001$, $OR = 3.29$, and no effect of condition, $b = -0.082$, $SE = 0.119$, $p = .49$, $OR = 0.92$. Crucially, we did not observe a significant interaction effect between facial trustworthiness and condition, $b = 0.107$, $SE = 0.143$, $p = .46$, $OR = 0.90$ (see Figure 3B). In line with the intuitive

accessibility account, but contrary to the subjective validity account, we found no evidence that participants relied less on facial trustworthiness when they could also rely on temptation, b = 1.139, SE = 0.094, $p < .001$, OR = 3.12, as opposed to when facial trustworthiness was the only available cue, b = 1.245, SE = 0.108, $p < .001$, OR = 3.47.

Response times. We again examined participants' response times to test how long they took to make trust decisions when relying on temptation or facial trustworthiness. Contrary to the results of Study 2, a *t*-test comparing response times between the temptation-only and face-only condition showed no evidence that participants in the face-only condition made faster decisions (M = 0.473, SD = 0.239) than participants in the temptation-only condition (M = 0.444, SD = 0.230), $t(880.47) = 1.85$, $p = .064$, $d = 0.12$.

To test the relationship between cue reliance and response times more directly, we regressed participants' average response times on their reliance on temptation, condition (coded -0.5 for temptation-only and 0.5 for face-and-temptation), and an interaction term between the two variables. This revealed a positive effect of condition, b = 0.046, SE = 0.015, $p = .002$, a positive effect of reliance on temptation, b = 0.040, SE = 0.007, $p < .001$, but no significant interaction effect between reliance on temptation and condition, b = 0.013, SE = 0.014, $p = .34$. Participants took longer to reach a decision when both temptation and facial trustworthiness were varied compared to when only temptation was varied. More importantly though, as in Study 2, we observed that the more participants relied on temptation, the longer they took to decide, and this effect did not significantly vary between the two conditions. Similarly, we regressed participants' average response times on their reliance on facial trustworthiness, condition (coded -0.5 for face-only and 0.5 for face-and-temptation), and an interaction term between the two variables. This revealed no effect of condition, b = 0.010, SE = 0.015, $p = .53$, no effect of reliance on facial trustworthiness, b = -0.008, SE = 0.007, $p = .25$, and no interaction effect between reliance on facial trustworthiness and condition, b = 0.001, SE = 0.014, $p = .94$. Thus, as in Study 2, we found no evidence that an increased reliance on facial trustworthiness was related to longer response times.

Discussion

In sum, participants still relied on facial trustworthiness in the presence of another (more subjectively valid) cue (i.e., economic payoffs). More importantly, the presence of facial cues led participants to rely somewhat *less* on the subjectively more valid cue. This pattern of results is in

line with the notion that people rely on facial trustworthiness because it is relatively quick and effortless.

Study 4

The results so far suggest that people favor relying on facial trustworthiness because it is intuitively accessible. Response time data from Study 2 and Study 3a-c provided support for this argument, as increased reliance on economic payoffs was related to longer decision times (while there was no relationship between reliance on facial cues and decision times). The goal of Study 4 was to provide experimental evidence for this claim by testing how reliance on economic payoffs and facial trustworthiness varies when participants make intuitive (vs. reflective) decisions. If reliance on economic payoffs requires cognitive reflection, we would expect an attenuated effect of payoff information when people make intuitive (vs. reflective) decisions.

Regarding reliance on facial trustworthiness, two predictions are plausible: People may override their trait impressions and rely more on economic information when making reflective as opposed to intuitive decisions, as they deem economic information to be the more valid cue. Alternatively, people may not be fully aware of how their decisions are influenced by their intuitive face judgments, or they may be unable to suppress or correct the influence of face judgments (T. D. Wilson & Brekke, 1994). From this perspective, additional reflection would not necessarily undermine the effect of face judgments on decisions.

Method

Participants. We recruited a sample of 962 U. S. American MTurk workers who participated in exchange for $2 each. Data from participants who failed an attention check at the end of the survey or who indicated having only a poor or basic English proficiency were excluded from analysis, leaving a final sample of 797 participants (47.55% female, M_{age} = 36.08, SD_{age} = 11.31).

Materials and procedure. The experiment was administered online. All participants learned the rules of the trust game and made a series of 24 hypothetical decisions, during which their interaction partners' temptation to betray and their facial trustworthiness were varied. We used the same twelve photos that were already used in Study 3b (six male and six female matched on attractiveness; three individuals with a happy and three with a neutral expression for each gender). Participants interacted twice with each individual—once when temptation was low

(0.2; 25% gain from betrayal) and once when temptation was high (0.60; 150% gain from betrayal). Participants were randomly assigned to one of two conditions.

In the intuition condition (n = 399), participants were prompted to follow their first instinct and make intuitive decisions (adapted from Newman, Gibb, & Thompson, 2017). We asked participants to reach each decision within five seconds. A timer counting backwards from five to zero was displayed on each decision page. Participants could still indicate a decision after the timer reached zero and the page only forwarded to the next trial once a decision was made. In the reflection condition (n = 398), participants were prompted to think carefully and make reflective decisions. Participants were informed that on each trial, they could only indicate a decision after ten seconds had passed. They were asked to take at least ten seconds to weigh all options and reflect on their decision. After they had made their decisions, participants were shown each face again and we asked them to rate how trustworthy they think the person in the photo is. We used the average trustworthiness rating of each face across all participants as our measure of facial trustworthiness.

Results

Descriptive statistics. Average trustworthiness ratings of the faces ranged from 36.35 to 74.24 (M = 55.25, SD = 13.19) and participants showed significant consensus in their ratings of the faces, ICC = .372, p < .001, 95% CI [.229, .631]. The average trust rate across all trials was 44.99% (SD = 25.44%) in the intuition condition and 37.32% (SD = 25.45%) in the reflection condition. The difference in trust rates between the two conditions was significant, $t(795) = 4.25$, p < .001, d = 0.30. A total of 91 participants (11.42%) never trusted whereas 26 participants (3.26%) always trusted. On average, participants in the intuition condition took 2.40 seconds to make a decision (SD = 4.04), whereas participants in the reflection condition took 13.66 seconds (SD = 0.98), $t(481.9) = 70.97$, p < .001, d = 5.03.

Reliance on temptation. We first tested the prediction that participants would rely less on payoff information when making intuitive (vs. reflective) decisions. We estimated a multilevel regression model with random intercepts per participant and random slopes for all predictors, in which we regressed participants' trust decisions on temptation, facial trustworthiness, condition (coded -0.5 for reflection and 0.5 for intuition), and an interaction term between temptation and condition. This yielded a negative effect of temptation, b = -0.670, SE = 0.059, p < .001, OR = 0.45, a positive effect of facial trustworthiness, b = 1.607, SE =

0.104, $p < .001$, $OR = 4.99$, and a positive effect of condition, $b = 0.547$, $SE = 0.136$, $p < .001$, $OR = 1.73$. Crucially, we observed a significant interaction effect between temptation and condition, $b = 1.185$, $SE = 0.115$, $p < .001$, $OR = 3.27$ (see Figure 4A). Participants relied less on temptation when making intuitive decisions, $b = -0.055$, $SE = 0.059$, $p = .35$, $OR = 0.95$, than when making reflective decisions, $b = -1.376$, $SE = 0.112$, $p < .001$, $OR = 0.25$.

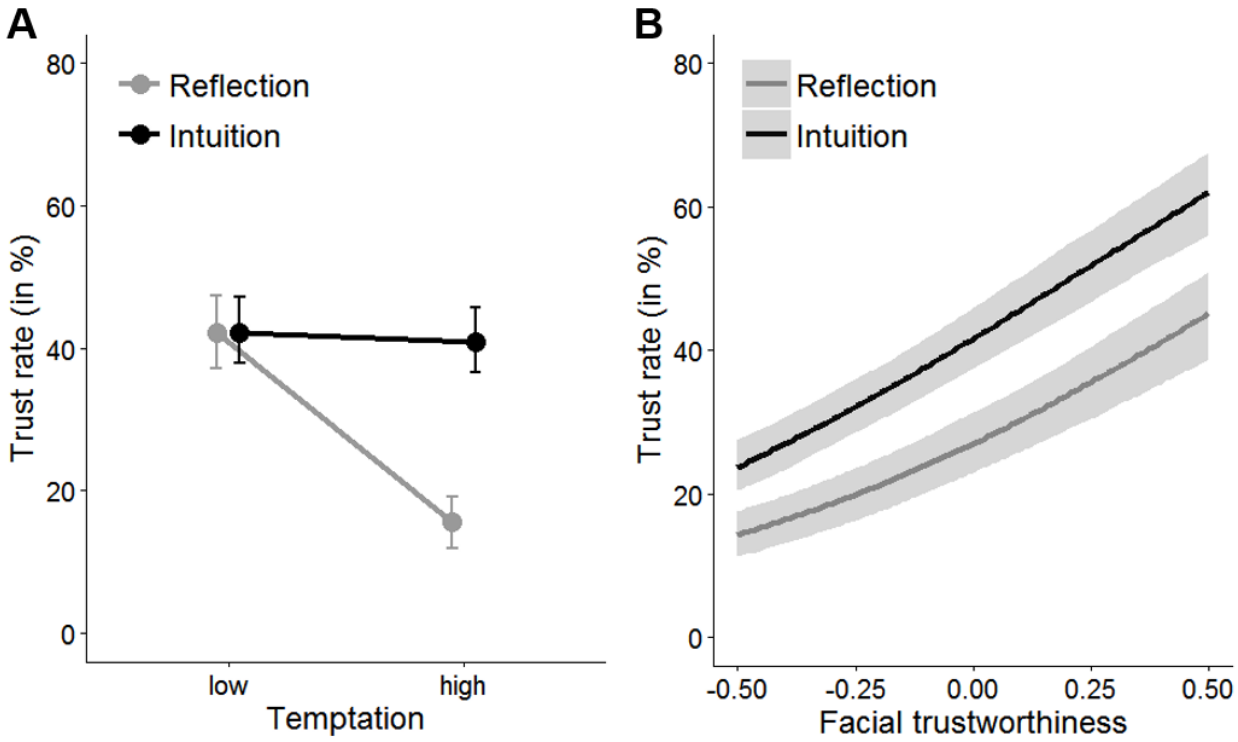

Figure 4. The effects of temptation and facial trustworthiness on trust rates when making intuitive versus reflective decisions (Study 4): (A) The effect of temptation on trust rates when participants made intuitive (vs. reflective) trust decisions. (B) The effect of facial trustworthiness on trust rates when participants made intuitive (vs. reflective) trust decisions. Values denote the predicted probability of trust derived from multilevel regression models.

Reliance on facial trustworthiness. Next, we tested how much participants relied on facial trustworthiness when making intuitive (vs. reflective) decisions. We again estimated a multilevel regression model, this time including an interaction term between facial trustworthiness and condition. This did not yield a significant interaction effect, $b = 0.237$, $SE = 0.218$, $p = .27$, $OR = 1.27$ (see Figure 4B). Participants relied on facial trustworthiness when

making intuitive decisions, $b = 1.660$, $SE = 0.153$, $p < .001$, $OR = 5.26$, and when making reflective decisions, $b = 1.596$, $SE = 0.142$, $p < .001$, $OR = 4.93$.[6]

Discussion

Results of Study 4 revealed that the influence of temptation, but not facial trustworthiness, on participants' trust decisions was reduced when participants made intuitive as opposed to reflective decisions. This finding is consistent with prior research suggesting that reliance on facial trustworthiness does not require cognitive reflection (Bonnefon et al., 2013; De Neys et al., 2017; Mieth, Bell, & Buchner, 2016). The results of the current study also shed light on the question of how the intuitive accessibility of face judgments causes their persistent effects on decision-making. Given enough time or motivation to reflect on their decisions, people might realize that relying on different information will lead to better decisions. However, it might also be the case that the effect of facial cues is more implicit: People could be unaware of how their decisions are influenced by intuitive face judgments, or they could be unable to suppress this influence (T. D. Wilson & Brekke, 1994). From this perspective, additional reflection would not necessarily undermine the effect of facial cues on decisions. The current results are in line with this latter view, as we did not find any evidence that the effect of facial trustworthiness is diminished when people made reflective as opposed to intuitive decisions.

General Discussion

We asked whether reliance on facial trustworthiness in social decision-making can be explained by beliefs in the diagnosticity of the cue (i.e., by subjective cue weighing) or by the fact that the cue is intuitively accessible which makes it relatively effortless to rely on it (i.e., by intuitive accessibility). Across six studies, we systematically tested how much people rely on facial trustworthiness and economic payoff information (i.e., the trustee's temptation to betray trust)—a cue that is perceived to be more valid but takes more effort to process. Across our studies, we find consistent support for the intuitive accessibility account.

In Study 1, we found that economic payoff information is seen as a more valid cue than facial information. Study 2 showed there was a positive correlation between reliance on economic payoffs and decision time, while no such relationship was found for reliance on facial

[6] Excluding 15 participants (1.88%), who indicated not having made intuitive or reflective decisions on all trials did not change the pattern of results.

trustworthiness. These findings replicated in Study 3a-c, with a substantially larger sample. To the extent that longer response times are an indicator for more effortful decision strategies (Bettman et al., 1990), these results suggest that it takes less cognitive effort to rely on facial trustworthiness rather than temptation. Thus, Study 1 identified economic payoff information as a cue that is seen as more valid, whereas Study 2 and Study 3a-c showed that economic payoff information is also more effortful to process than facial trustworthiness.

In Study 3a-c, we tested to what extent people rely on economic payoff information and facial trustworthiness when both cues are available simultaneously. Results showed that people rely less on a subjectively more valid cue (i.e., temptation) when facial trustworthiness is available as well. However, people do not rely less on facial trustworthiness when temptation is simultaneously manipulated. This pattern of results is in line with the argument that persistent reliance on facial trustworthiness is driven by the intuitive accessibility of the cue. If cognitive reflection is required for reliance on economic payoff information, but not for reliance on facial cues, then we would expect that restricting reflection during decision-making only decreases the influence of payoff information, but not the influence of facial cues. This prediction was explicitly tested in Study 4. This final study confirmed that payoff information had a diminished effect on trust when people made intuitive (vs. reflective) decisions. Taken together, our data suggest that the persistent reliance on facial trustworthiness can be better explained by the intuitively accessibility of the cue, which makes reliance on it relatively effortless, rather than by the belief that it is a particularly valid cue.

Decision effort and trust

The present results converge with prior research on the processing of trait information from faces. A host of studies has demonstrated that trustworthiness impressions from faces are formed in a particularly efficient manner (Bonnefon et al., 2013; Klapper et al., 2016; Willis & Todorov, 2006). We extend these findings by showing that reliance on economic payoff information requires more effort than reliance on facial cues. Similar results were observed when we manipulated how much people could reflect on their decisions. Intuitive decisions were less reliant on economic payoff information, but the manipulation of reflection time had no effect on reliance on facial cues. Our interpretation that reliance on facial trustworthiness is relatively effortless is also in line with studies showing that subjecting participants to a cognitive load

manipulation does not impair their use of facial cues when making trust decisions (Bonnefon et al., 2013; De Neys et al., 2017; Mieth et al., 2016).

In Study 2 and Study 3a-c, we measured response times as a proxy for the cognitive effort needed to make a decision (Bettman et al., 1990). Previous investigations in the domain of social decision-making have highlighted that other factors such as decision conflict can also drive response times (Evans, Dillon, & Rand, 2015; Krajbich et al., 2015). Perhaps, decision-makers felt more conflicted when relying on temptation, which increased their decision times. According to this view, we would also expect participants to be less confident when relying on temptation as decision conflict is associated with decreased confidence (De Neys et al., 2011; Zakay, 1985). Yet, across our studies, we found no evidence that participants reported lower levels of confidence in the presence of temptation as opposed to facial trustworthiness, suggesting that decision conflict was not driving our results.

Our results also fit within broader frameworks of how people decide whom to trust. With regard to the question which cues people rely on in trust situations, Evans and Krueger (2016) suggest in their model of bounded prospection that people approach trust decisions egocentrically: People focus on their potential costs and benefits while neglecting the probability that these outcomes will occur as assessing an interaction partner's trustworthiness requires perspective-taking and thus cognitive effort (S. Lin, Keysar, & Epley, 2010). We go beyond contrasting cues pertaining to the self with cues pertaining to the situation (Thielmann & Hilbig, 2015) or the interaction partner (Evans & Krueger, 2011) and provide evidence for a more general claim that the extent to which people rely on information in trust decisions with multiple cues is determined by the ease with which the information is processed.

Limitations and future research

We provided evidence that the persistent reliance on facial trustworthiness can at least partly be explained by the intuitive accessibility of face judgments. We demonstrated this by showing that people favor relying on facial trustworthiness over temptation even though the latter is seen as a more valid cue. More specifically, we found that people did not rely less on facial trustworthiness when they could also rely on temptation. This finding stands in contrast to previous studies showing that people discount facial cues when other information is available (Graham, Harvey, & Puri, 2017; Rezlescu et al., 2012; Yu, Saleem, & Gonzalez, 2014). For example, Rezlescu and colleagues (2012) showed that, the influence of facial trustworthiness on

trust decisions was reduced, but not eliminated, when information on a partner's past behavior was available. Maybe the subjective validity of temptation is still comparatively low and a stronger cue would have reduced or even eliminated reliance on facial cues in our studies. Arguably, people may prioritize past behavior over facial trustworthiness because past behavior is both easy to process and subjectively valid. To date, there is relatively little work that examines how people prioritize different types of cues in dilemmas of trust; future studies need to test how reliance on facial trustworthiness changes in the presence of different cues that vary in both subjective validity and ease of accessibility.

In addition, more research is needed to understand how the efficient processing of faces influences cue selection and cue weighing. Following the fast accessibility of face judgments, do people ignore other cues altogether? Our results speak against such a strong version of one-reason decision-making (Gigerenzer et al., 2008) as the effect of temptation was still apparent even when decision-makers could have relied solely on facial trustworthiness. Individual differences in rational versus heuristic processing (Epstein, Pacini, Denes-Raj, & Heier, 1996) or need for cognition (Cacioppo & Petty, 1982) might help explain why some individuals reach a decision after face judgments are accessible whereas other consider additional cues. People who score higher on heuristic processing may rely more on facial trustworthiness even when other cues are available.

Alternatively, it could be the case that our participants did in fact consider all available cues but that the efficient processing of faces influenced how the cues were weighed. Shah (2007) has argued that efficiently processed cues are weighed more heavily. Since participants only read a description of the face cue and did not experience the efficient processing of it in Study 1, our measure of subjective cue validity might have underestimated how valid participants perceived the cue to be *while* making decisions. However, Dimov and Link (2017) showed that processing efficiency does not affect cue weights, but rather the order in which cues are considered. In the context of our study, it is thus conceivable that the discounting of temptation was due to a sequential process where fast facial trustworthiness judgments anchored participants' response and subsequent adjustment on the basis of temptation was insufficient (Tamir & Mitchell, 2012). This view holds that the primary processing of faces influences the weighing of subsequently processed cues. Future research could test whether a serial

presentation of cues, where participants first learn about temptation and then see the face of their interaction partner, eliminates the discounting of temptation in favor of facial trustworthiness.[7]

Implications for debiasing interventions

The primary goal of our current investigation was to shed light on the process giving rise to the face bias. In addition, our findings may also be used to inform the design of interventions aimed at curbing the bias. One previously proposed recommendation is to inform decision-makers about the poor accuracy of their impressions (Porter et al., 2010). Our results suggest that this approach might not be sufficient as the face bias does not stem from a conscious (albeit subjective) weighing of cues (Dawes et al., 1989; T. D. Wilson & Brekke, 1994). Another idea would be to increase the importance of the decision in order to motivate people to process the available information more deeply (Chaiken, 1980; Chaiken & Maheswaran, 1994) or to reduce processing constraints (e.g., time pressure) or decision complexity (Gigerenzer & Goldstein, 1996; Payne et al., 1988). Both approaches might lead to an attenuation of the face bias but we are skeptical that it will completely eliminate reliance on facial cues. Previous studies have shown that the influence of facial trustworthiness persists for extremely important decisions, such as legal sentencing (Blair, Judd, & Fallman, 2004; Eberhardt, Davies, Purdie-Vaughns, & Johnson, 2006; J. P. Wilson & Rule, 2015) or voting (Olivola & Todorov, 2010a), as well as for relatively simple, self-paced decisions in the lab (Mieth et al., 2016; Rezlescu et al., 2012). As a further case in point, Study 4 showed that, while making people reflect on their decision increased their reliance on temptation, we found no evidence that it reduced the influence of facial trustworthiness.

Our findings suggest that interventions that disrupt the primary and efficient processing of faces might hold more potential for success. Inverting photos or misaligning its parts might disrupt the efficient processing of faces while still enabling the identification of an individual

[7] An anonymous reviewer raised the possibility that people might simply be reluctant to report that they see faces as valid cues. While more research is needed to measure how accurate people think their face judgments are, the available evidence suggests that people do think and report that faces contain information about a person's personality (Hassin & Trope, 2000; Suzuki et al., 2017). For example, in a survey by Hassin and Trope (2000), they found that 75% of respondents agreed that at least some traits can be read from a person's facial appearance.

(Bindemann, Burton, Hooge, Jenkins, & de Haan, 2005; Todorov, Loehr, & Oosterhof, 2010), but requesting photos to be displayed upside-down in case files or on websites such as Airbnb is not realistic. However, in many situations, decision-makers could be asked to reach a decision in absence of facial cues first—ideally based on valid information. In the next step, a photo is displayed while the decision can still be revised. With this setup, initial decisions should be less biased. Any incorporation of face judgments into the decision-making process would now require a conscious revision of the decision, which is less likely if more diagnostic cues are also available. Future studies should test interventions targeting the efficient processing of faces and compare their effectiveness to simply teaching decision-makers about the biasing potential of faces.

Conclusion

Despite their poor predictive validity, face judgments influence decisions ranging from the relatively trivial choice of which apartment to rent (which might determine the fate of the next vacation; Ert et al., 2016), to the selection of a CEO (which might determine the fate of a company; Gomulya et al., 2016; Graham, Harvey, & Puri, 2017; Stoker, Garretsen, & Spreeuwers, 2016), to the sentencing of a criminal (which might determine the fate of a human being; Blair, Judd, & Fallman, 2004; Eberhardt et al., 2006; J. P. Wilson & Rule, 2015). It may thus not be surprising that researchers have called for a more nuanced understanding of why people persistently rely on face judgments (Olivola et al., 2014). We contribute to this debate by showing that peoples' reliance on face judgments is better explained by the intuitive accessibility of the cue, which make reliance on it relatively effortless, rather than by beliefs that facial cues are particularly valid. We recommend that future attempts at designing interventions focus on disrupting the primary and efficient processing of faces to attenuate the effects of face judgments on social decisions.

References

Aristotle. (1936). Physiognomics. In T. E. Page, E. Capps, W. H. D. Rouse, A. Post, & E. H. Warmington (Eds.), *Minor Works* (1955th ed.). Cambridge, Massachusetts: Harvard University Press.

Bacharach, M., & Gambetta, D. (2001). Trust in signs. In K. S. Cook (Ed.), *Trust in Society* (pp. 148–184). Russell Sage Foundation.

Berg, J., Dickhaut, K., & McCabe, K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior*, *10*(1), 122–142.

Berry, D. S., & Zebrowitz-McArthur, L. A. (1988). What's in a face? Facial maturity and the attribution of legal responsibility. *Personality & Social Psychology Bulletin*, *14*(1), 23–33.

Bettman, J. R., Johnson, E. J., & Payne, J. W. (1990). A componential analysis of cognitive effort in choice. *Organizational Behavior and Human Decision Processes*, *45*, 111–139.

Bindemann, M., Burton, A. M., Hooge, I. T. C., Jenkins, R., & de Haan, E. H. F. (2005). Faces retain attention. *Psychonomic Bulletin & Review*, *12*(6), 1048–1053. https://doi.org/10.3758/BF03206442

Blair, I. V, Judd, C. M., & Fallman, J. L. (2004). The automaticity of race and Afrocentric facial features in social judgments. *Journal of Personality and Social Psychology*, *87*(6), 763–778. https://doi.org/10.1037/0022-3514.87.6.763

Bonnefon, J. F., Hopfensitz, A., & De Neys, W. (2013). The modular nature of trustworthiness detection. *Journal of Experimental Psychology: General*, *142*(1), 143–150. https://doi.org/10.1037/a0028930

Bonnefon, J. F., Hopfensitz, A., & De Neys, W. (2015). Face-ism and kernels of truth in facial inferences. *Trends in Cognitive Sciences*, *19*(8), 421–422. https://doi.org/10.1016/j.tics.2015.05.002

Bonnefon, J. F., Hopfensitz, A., & De Neys, W. (2017). Can we detect cooperators by looking at their face? *Current Directions in Psychological Science*, *26*(3), 276–281.

Brunswik, E. (1956). *Perception and the representative design of psychological experiments*. Berkeley: Univer. California Press.

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, *42*(1), 116–131.

Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology*, *39*(5), 752–766.

Chaiken, S., & Maheswaran, D. (1994). Heuristic processing can bias systematic processing: Effects of sourcecredibility, argument ambiguity, and task importance on attitude judgment. *Journal of Personality and Social Psychology*, *66*(3), 460–473.

Costa-Gomes, M. A., Huck, S., & Weizsäcker, G. (2014). Beliefs and actions in the trust game: Creating instrumental variables to estimate the causal effect. *Games and Economic Behavior*, *88*, 298–309. https://doi.org/10.1016/j.geb.2014.10.006

Dawes, R. M. (1980). Social dilemmas. *Annual Review of Psychology*, *31*, 169–193. https://doi.org/10.1146/annurev.ps.31.020180.001125

Dawes, R. M., Faust, D., & Meehl, P. E. (1989). Clinical versus actuarial judgment. *Science*, *243*(4899), 1668–1674. https://doi.org/10.1126/science.2648573

De Neys, W., Cromheeke, S., & Osman, M. (2011). Biased but in doubt: Conflict and decision confidence. *PLoS ONE*, *6*(1). https://doi.org/10.1371/journal.pone.0015954

De Neys, W., Hopfensitz, A., & Bonnefon, J.-F. (2017). Split-second trustworthiness detection from faces in an economic game. *Experimental Psychology*, *64*, 231–239.

Delacre, M., Lakens, D., & Leys, C. (2017). Why psychologists should by default use Welch's t-test instead of Student's t- test. *International Review of Social Psychology*, *30*(1), 92–101. https://doi.org/http://doi.org/10.5334/irsp.82

Dimov, C. M., & Link, D. (2017). Do people order cues by retrieval fluency when making probabilistic inferences? *Journal of Behavioral Decision Making*. https://doi.org/10.1002/bdm.2002

Duarte, J., Siegel, S., & Young, L. (2012). Trust and credit: The role of appearance in peer-to-peer lending. *Review of Financial Studies*, *25*(8), 2455–2483. https://doi.org/10.1093/rfs/hhs071

Dzhelyova, M., Perrett, D. I., & Jentzsch, I. (2012). Temporal dynamics of trustworthiness perception. *Brain Research, 1435*, 81–90. https://doi.org/10.1016/j.brainres.2011.11.043

Eberhardt, J. L., Davies, P. G., Purdie-Vaughns, V. J., & Johnson, S. L. (2006). Looking deathworthy. *Psychological Science*, *17*(5), 383–386. https://doi.org/10.1111/j.1467-9280.2006.01716.x

Efferson, C., & Vogt, S. (2013). Viewing men's faces does not lead to accurate predictions of trustworthiness. *Scientific Reports*, *3*, 1047. https://doi.org/10.1038/srep01047

Epstein, S., Pacini, R., Denes-Raj, V., & Heier, H. (1996). Individual differences in intuitive-experiential and analytical-rational thinking styles. *Journal of Personality and Social Psychology*, *71*(2), 390–405. https://doi.org/10.1037/0022-3514.71.2.390

Ert, E., Fleischer, A., & Magen, N. (2016). Trust and reputation in the sharing economy: The role of personal photos in Airbnb. *Tourism Management*, *55*, 62–73. https://doi.org/10.1016/j.tourman.2016.01.013

Evans, A. M., Dillon, K. D., & Rand, D. G. (2015). Fast but not intuitive, slow but not reflective: Decision conflict drives reaction times in social dilemmas. *Journal of Experimental Psychology: General*, *144*(5), 951–966. https://doi.org/10.1037/xge0000107

Evans, A. M., & Krueger, J. I. (2011). Elements of trust: Risk and perspective-taking. *Journal of Experimental Social Psychology*, *47*(1), 171–177. https://doi.org/10.1016/j.jesp.2010.08.007

Evans, A. M., & Krueger, J. I. (2014). Outcomes and expectations in dilemmas of trust. *Judgment and Decision Making*, *9*(2), 90–103.

Evans, A. M., & Krueger, J. I. (2016). Bounded prospection in dilemmas of trust and reciprocity. *Reviews of General Psychology*, *20*(1), 1–5. https://doi.org/10.1007/s13398-014-0173-7.2

Evans, A. M., & Rand, D. G. (2019). Cooperation and decision time. *Current Opinion in Psychology*, *26*, 67–71. https://doi.org/10.1016/j.copsyc.2018.05.007

Evans, A. M., & van de Calseyde, P. P. F. M. (2017). The effects of observed decision time on expectations of extremity and cooperation. *Journal of Experimental Social Psychology*, *68*, 50–59. https://doi.org/10.1016/j.jesp.2016.05.009

Ewing, L., Caulfield, F., Read, A., & Rhodes, G. (2014). Perceived trustworthiness of faces drives trust behaviour in children. *Developmental Science*, *2*, 327–334. https://doi.org/10.1111/desc.12218

Freeman, J. B., & Johnson, K. L. (2016). More than meets the eye: Split-second social perception. *Trends in Cognitive Sciences*, *20*(5), 362–374. https://doi.org/10.1016/j.tics.2016.03.003

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, *103*(4), 650–69. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8888650

Gigerenzer, G., Hertwig, R., & Pachur, T. (2011). *Heuristics: The foundations of adaptive behavior*. New York, NY: Oxford University Press.

Gigerenzer, G., Martignon, L., Hoffrage, U., Rieskamp, J., Czerlinski, J., & Goldstein, D. G. (2008). One-reason decision making. In C. R. Plott & V. L. Smith (Eds.), *Handbook of Experimental Economics* (Vol. 1, pp. 1004–1017). Amsterdam: North-Holland. https://doi.org/10.1016/S1574-0722(07)00108-4

Gomulya, D., Wong, E. M., Ormiston, M. E., & Boeker, W. (2017). The role of facial appearance on CEO selection after firm misconduct. *Journal of Applied Psychology*, *102*(4), 617–635. https://doi.org/http://dx.doi.org/10.1037/apl0000172

Graham, J. R., Harvey, C. R., & Puri, M. (2017). A corporate beauty contest. *Management Science*, *63*(9), 3044–3056. https://doi.org/10.1287/mnsc.2016.2484

Hammond, K. R., Hursch, C. J., & Todd, F. J. (1964). Analyzing the components of clinical inference. *Psychological Review*, *71*(6), 438–456. https://doi.org/10.1037/h0040736

Hassin, R., & Trope, Y. (2000). Facing faces: Studies on the cognitive aspects of physiognomy. *Journal of Personality and Social Psychology*, *78*(5), 837–852. https://doi.org/10.1037//0022-3514.78.5.837

Kelley, H. H., Holmes, J. G., Kerr, Norbert, L., Reis, H. T., Rusbult, C. E., & Van Lange, P. A. M. (2003). *An Atlas of Interpersonal Situations*. Cambridge University Press.

Klapper, A., Dotsch, R., van Rooij, I., & Wigboldus, D. H. J. (2016). Do we spontaneously form stable trustworthiness impressions from facial appearance? *Journal of Personality and Social Psychology*, *111*(5), 655–664. https://doi.org/10.1037/pspa0000062

Krajbich, I., Bartling, B., Hare, T., & Fehr, E. (2015). Rethinking fast and slow based on a critique of reaction-time reverse inference. *Nature Communications*, *6*. https://doi.org/10.1038/ncomms8455

Krumhuber, E. G., Manstead, A. S. R. R., Cosker, D., Marshall, D., Rosin, P. L., & Kappas, A. (2007). Facial dynamics as indicators of trustworthiness and cooperative behavior. *Emotion*, *7*(4), 730–5. https://doi.org/10.1037/1528-3542.7.4.730

Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H. J., Hawk, S. T., & van Knippenberg, A. (2010). Presentation and validation of the Radboud Faces Database. *Cognition & Emotion*, *24*(8), 1377–1388. https://doi.org/10.1080/02699930903485076

Lavater, J. C. (1775). *Essays on Physiognomy: Designed to Promote the Knowledge and the Love of Mankind*. London: William Tegg and Co.

Li, T., Liu, X., Pan, J., & Zhou, G. (2017). The interactive effect of facial appearance and behavior statement on trust belief and trust behavior. *Personality and Individual Differences*, *117*, 60–65. https://doi.org/10.1016/j.paid.2017.05.038

Lin, C., Adolphs, R., & Alvarez, R. M. (2018). Inferring whether officials are corruptible from looking at their faces. *Psychological Science*, 1–7. https://doi.org/10.1177/0956797618788882

Lin, S., Keysar, B., & Epley, N. (2010). Reflexively mindblind: Using theory of mind to interpret behavior requires effortful attention. *Journal of Experimental Social Psychology*, *46*(3), 551–556. https://doi.org/10.1016/j.jesp.2009.12.019

McCullough, M. E., & Reed, L. I. (2016). What the face communicates: Clearing the conceptual ground. *Current Opinion in Psychology*, *7*, 110–114. https://doi.org/10.1016/j.copsyc.2015.08.023

Mieth, L., Bell, R., & Buchner, A. (2016). Cognitive load does not affect the behavioral and cognitive foundations of social cooperation. *Frontiers in Psychology*, *7*, 1–14. https://doi.org/10.3389/fpsyg.2016.01312

Newman, I. R., Gibb, M., & Thompson, V. A. (2017). Rule-based reasoning is fast and belief-based reasoning can be slow: Challenging current explanations of belief-bias and base-rate neglect. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *43*(7), 1154–1170.

Olivola, C. Y., Funk, F., & Todorov, A. (2014). Social attributions from faces bias human choices. *Trends in Cognitive Sciences*, *18*(11), 566–70. https://doi.org/10.1016/j.tics.2014.09.007

Olivola, C. Y., & Todorov, A. (2010a). Elected in 100 milliseconds: Appearance-based trait inferences and voting. *Journal of Nonverbal Behavior*, *34*(2), 83–110. https://doi.org/10.1007/s10919-009-0082-1

Olivola, C. Y., & Todorov, A. (2010b). Fooled by first impressions? Reexamining the diagnostic value of appearance-based inferences. *Journal of Experimental Social Psychology*, *46*(2), 315–324. https://doi.org/10.1016/j.jesp.2009.12.002

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences*, *105*(32), 11087–11092. https://doi.org/10.1073/pnas.0805664105

Paolacci, G., & Chandler, J. (2014). Inside the Turk: Understanding Mechanical Turk as a participant pool. *Current Directions in Psychological Science*, *23*(3), 184–188. https://doi.org/10.1177/0963721414531598

Payne, J., Bettman, J. R., & Johnson, E. J. (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *14*(3), 522–534.

Porter, S., ten Brinke, L., & Gustaw, C. (2010). Dangerous decisions: The impact of first impressions of trustworthiness on the evaluation of legal evidence and defendant culpability. *Psychology, Crime & Law*, *16*(6), 477–491. https://doi.org/10.1080/10683160902926141

Rezlescu, C., Duchaine, B., Olivola, C. Y., & Chater, N. (2012). Unfakeable facial configurations affect strategic choices in trust games with or without information about past behavior. *PLoS ONE*, *7*(3), e34293. https://doi.org/10.1371/journal.pone.0034293

Ro, T., Russell, C., & Lavie, N. (2001). Changing faces: a detection advantage in the flicker paradigm. *Psychological Science*, *12*(1), 94–99. https://doi.org/10.1111/1467-9280.00317

Rule, N. O., Krendl, A. C., Ivcevic, Z., & Ambady, N. (2013). Accuracy and consensus in judgments of trustworthiness from faces: Behavioral and neural correlates. *Journal of Personality and Social Psychology*, *104*(3), 409–26. https://doi.org/10.1037/a0031050

Said, C. P., Sebe, N., & Todorov, A. (2009). Structural resemblance to emotional expressions predicts evaluation of emotionally neutral faces. *Emotion*, *9*(2), 260–264. https://doi.org/10.1037/a0014681

Scharlemann, J. P. ., Eckel, C. C., Kacelnik, A., & Wilson, R. K. (2001). The value of a smile: Game theory with a human face. *Journal of Economic Psychology*, *22*(5), 617–640. https://doi.org/10.1016/S0167-4870(01)00059-9

Shah, A. K. (2007). Easy does it: The role of fluency in cue weighting. *Judgment and Decision Making*, *2*(6), 371–379. https://doi.org/10.1037/e722852011-015

Shah, A. K., & Oppenheimer, D. M. (2008). Heuristics made easy: an effort-reduction framework. *Psychological Bulletin*, *134*(2), 207–222. https://doi.org/10.1037/0033-2909.134.2.207

Simmons, J. P., & Nelson, L. D. (2006). Intuitive confidence: choosing between intuitive and nonintuitive alternatives. *Journal of Experimental Psychology: General*, *135*(3), 409–428. https://doi.org/10.1037/0096-3445.135.3.409

Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, *69*(1), 99–118. https://doi.org/10.2307/1884852

Slepian, M. L., & Ames, D. R. (2015). Internalized impressions: The link between apparent facial trustworthiness and deceptive behavior Is mediated by targets' expectations of how they will be judged. *Psychological Science*, 1–7. https://doi.org/10.1177/0956797615594897

Slovic, P., & Lichtenstein, S. (1971). Comparison of Bayesian and regression approaches to the study of information processing in judgement. *Organizational Behavior and Human Performance*, *6*, 649–744.

Snijders, C. C. P. C., & Keren, G. G. (1999). Determinants of trust. In D. V Budescu, I. Erev, & R. Zwick (Eds.), *Games and human behavior: Essays in honor of Amnon Rapoport* (pp. 355–385). Mawah, NJ: Erlbaum.

Stirrat, M., & Perrett, D. I. (2010). Valid facial cues to cooperation and trust: Male facial width and trustworthiness. *Psychological Science*, *21*(3), 349–354. https://doi.org/10.1177/0956797610362647

Stoker, J. I., Garretsen, H., & Spreeuwers, L. J. (2016). The facial appearance of CEOs: Faces signal selection but not performance. *PloS One*, *11*(7), e0159950. https://doi.org/10.1371/journal.pone.0159950

Sutherland, C. A. M., Young, A. W., & Rhodes, G. (2016). Facial first impressions from another angle: How social judgements are influenced by changeable and invariant facial properties. *British Journal of Psychology*, 1–19. https://doi.org/10.1111/bjop.12206

Suzuki, A., Tsukamoto, S., & Takahashi, Y. (2017). Faces tell everything in a just and biologically determined world. *Social Psychological and Personality Science*, 194855061773461. https://doi.org/10.1177/1948550617734616

Tamir, D. I., & Mitchell, J. P. (2012). Anchoring and adjustment during social inferences. *Journal of Experimental Psychology: General*, *142*(1), 151–162. https://doi.org/10.1037/a0028232

Theeuwes, J., & Van der Stigchel, S. (2006). Faces capture attention: Evidence from inhibition of return. *Visual Cognition*, *13*(6), 657–665. https://doi.org/10.1080/13506280500410949

Thielmann, I., & Hilbig, B. E. (2014). Trust in me, trust in you: A social projection account of the link between personality, cooperativeness, and trustworthiness expectations. *Journal of Research in Personality*, *50*(1), 61–65. https://doi.org/10.1016/j.jrp.2014.03.006

Thielmann, I., & Hilbig, B. E. (2015). Trust: An integrative review from a person-situation perspective. *Review of General Psychology*, *19*(3), 249–277. https://doi.org/10.1037/gpr0000046

Todorov, A., Funk, F., & Olivola, C. Y. (2015). Response to Bonnefon et al.: Limited 'kernels of truth' in facial inferences. *Trends in Cognitive Sciences*, *19*(8), 422. https://doi.org/10.1016/j.tics.2015.05.002

Todorov, A., Loehr, V., & Oosterhof, N. N. (2010). The obligatory nature of holistic processing of faces in social judgments. *Perception*, *39*(4), 514–532. https://doi.org/10.1068/p6501

Todorov, A., Olivola, C. Y., Dotsch, R., & Mende-Siedlecki, P. (2015). Social attributions from faces: Determinants, consequences, accuracy, and functional significance. *Annual Review of Psychology*, *66*(1), 519–545. https://doi.org/10.1146/annurev-psych-113011-143831

Todorov, A., & Porter, J. M. (2014). Misleading first impressions: Different for different facial images of the same person. *Psychological Science*, *25*(7), 1404–1417. https://doi.org/10.1177/0956797614532474

Tognetti, A., Berticat, C., Raymond, M., & Faurie, C. (2013). Is cooperativeness readable in static facial features? An inter-cultural approach. *Evolution and Human Behavior*, *34*(6), 427–432. https://doi.org/10.1016/j.evolhumbehav.2013.08.002

van't Wout, M., & Sanfey, A. G. (2008). Friend or foe: The effect of implicit trustworthiness judgments in social decision-making. *Cognition*, *108*, 796–803. https://doi.org/10.1016/j.cognition.2008.07.002

Vogt, S., Efferson, C., & Fehr, E. (2013). Can we see inside? Predicting strategic behavior given limited information. *Evolution and Human Behavior*, *34*(4), 258–264. https://doi.org/10.1016/j.evolhumbehav.2013.03.003

Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after a 100-ms exposure to a face. *Psychological Science, 17*(7), 592–598. https://doi.org/10.1111/j.1467-9280.2006.01750.x

Wilson, J. P., & Rule, N. O. (2015). Facial trustworthiness predicts extreme criminal-sentencing outcomes. *Psychological Science, 26*(8), 1325–1331. https://doi.org/10.1177/0956797615590992

Wilson, J. P., & Rule, N. O. (2016). Hypothetical sentencing decisions are associated with actual capital punishment outcomes: The role of facial trustworthiness. *Social Psychological and Personality Science, 7*(4), 331–338. https://doi.org/10.1177/1948550615624142

Wilson, R. K., & Eckel, C. C. (2006). Judging a book by its cover: Beauty and expectations in the trust game. *Political Research Quarterly*, *59*(2), 189–202. https://doi.org/10.1177/106591290605900202

Wilson, T. D., & Brekke, N. (1994). Mental contamination and mental correction: Unwanted influences on judgments and evaluations. *Psychological Bulletin, 116*(1), 117–142. https://doi.org/10.1037/0033-2909.116.1.117

Yu, M., Saleem, M., & Gonzalez, C. (2014). Developing trust: First impressions and experience. *Journal of Economic Psychology*, *43*, 16–29. https://doi.org/10.1016/j.joep.2014.04.004

Zakay, D. (1985). Post-decisional confidence and conflict experienced in a choice process. *Acta Psychologica, 58*, 75–80.